Report No. 67-56
October 1967

CONTROLLABILITY OF DISTRIBUTED PARAMETER SYSTEMS

by

Charles John Herget

Department of Engineering
University of California
Los Angeles

1967

FOREWORD

The research described in this report, "Controllability of Distributed Parameter Systems," Number 67-56, by C.J. Herget, was carried out under the direction of C.T. Leondes, in the Department of Engineering, University of California, Los Angeles.

This research was supported by the National Aeronautics and Space Administration, Grant NAS8-21019, Adaptive Control for Flexible Boosters, and the NASA Grant NsG 237-62 to the Institute of Geophysics and Interplanetary Physics of the University.

This report is based on a Doctor of Philosophy dissertation submitted by the author.

TABLE OF CONTENTS

ABSTRACT

The subject of this report is the controllability of distributed parameter systems. Two closely related topics also covered are minimum energy control systems and the reachable set of states with a norm constraint on the control.

A summary of the techniques which are applicable to the solution of control systems problems is given. The eigenvalue-eigenfunction expansion method for the solution of homogeneous boundary value problems is used. Problems in which the control appears at the boundary are treated by converting the non-homogeneous b.v.p. to an equivalent homogeneous b.v.p. by introducing generalized functions.

The generalization of the concept of controllability of finite dimensional systems to infinite dimensional systems is given. The pseudo-inverse of a linear operator is defined which is a generalization of that of a matrix for finite dimensional spaces. The pseudo-inverse is then used to obtain minimum energy control for distributed parameter systems. It is shown that this generalization includes results for finite dimensional systems which are available. In this case the solution of the minimum energy control problem involves finding the pseudo-inverse of a matrix. In the infinite dimensional problem, it is necessary to solve for the eigenvalues and eigenfunctions of an integral operator.

The necessary and sufficient conditions for the states which are reachable when the control is required to satisfy a norm constraint are given. The conditions are obtained by an application of the moment problem to distributed parameter systems. These results are then used to obtain conditions for complete controllability.

Applications of the material are made to specific examples of control systems problems.

CHAPTER 1

INTRODUCTION

The area of research of this dissertation is in the control of distributed parameter systems.

Until a few years ago, all of the emphasis of control systems was placed on systems described by ordinary differential equations. A well formulated theory has developed around these systems, particularly in the case of linear systems. In an attempt to obtain more generality, control systems engineers naturally attempted to obtain results for other systems which required control. One of the important class of systems is that of the type described by partial differential equations. These systems are described by the more illustrative engineering term, distributed parameter systems.

One of the most comprehensive articles appearing so far on this topic is that by Wang in [1]. He covers the wide range of problems of interest to control systems engineers. He has taken many of the problems which have been solved for finite dimensional systems and reformulated them as problems applied to distributed parameter systems. A major part of his work, as with most papers appearing on distributed parameter systems, is on optimum control, that is, finding a control which will minimize a specified cost functional.

The topic of controllability has found a great deal of interest in the finite dimensional theory and is a significant topic for research under more general circumstances. Very little work has appeared in the engineering literature on the subject of controllability of distributed parameter systems, and this will be the subject of this dissertation. Two closely related topics to controllability are also covered. These are minimum energy control systems and the reachable set of states with a norm constraint on the control. The only related work on distributed parameter known to the author appears in Wang [1] and Brogan [2]. Both touch the subject only briefly. More will be said concerning their results in Chapter 3. Abstract results on the theoretical aspects of controllability have been obtained by Fattorini [3, 4]. His problems are set in an abstract Banach space as is most of the work on the mathematical theory of control which has appeared recently, for example, by Balakrishnan [5]. Russell [6] has presented some material on controllability of distributed parameter systems which is to be published soon.

The term controllability was introduced by Kalman for finite dimensional control systems around 1960 and has become a fundamental concept in the presently developing field of systems theory. A general summary of the results available for finite dimensional linear control systems can be found in Kalman, et. al. [7]. The minimum energy control of systems is also found in [7] where use is made of the pseudo-inverse of a matrix.

Hsieh [8] approached the minimum effort control system problem for finite dimensional systems by setting the problem in a Hilbert space and found convenient results through the use of functional analysis.

Antosiewicz [9] found conditions for controllability based on geometric ideas in Banach spaces and obtained a result very similar to that which was given by Banach in [10] as conditions for the solution of the moment problem. Antosiewicz, however, restricted his results to systems whose states are finite dimensional.

Kreindler [11] obtained results on the set of reachable states with a norm constraint on the control for finite dimensional linear systems.

Numerous other papers have appeared on the control of distributed parameter systems. For the most part, they are concerned with optimum control and so they will not be mentioned here.

A brief review of the work in the remaining chapters of this dissertation will now be given.

Chapter 2 is titled "Partial Differential Equations in Control Systems Applications." There are many books on partial differential equations; however, there is no particular reference suitably oriented to control systems applications. Thus Chapter 2 is a summary of the techniques which have been applied elsewhere which seem appropriate to the topic of this dissertation. The eigenvalue-eigenfunction method of solving non-homogeneous partial differential equations with homogeneous boundary conditions is presented. The solution of problems where the control appears at the boundary is treated next. This situation leads to non-homogeneous boundary value problems. Several methods are available for handling this type of problem. The method presented in Chapter 2 follows that of Friedman [12] where the non-homogeneous b.v.p. is changed to an equivalent non-homogeneous equation with homogeneous boundary conditions by introducing generalized functions.

Chapter 3 is titled "Controllability." This chapter contains the generalization of the concept of controllability to infinite dimensional systems. The pseudo-inverse of a linear operator is defined which is a generalization of that of a matrix for finite dimensional spaces. The pseudo-inverse is then used to obtain minimum energy control for distributed parameter systems.

Chapter 4 is titled "Reachable States." The necessary and sufficient conditions on the set of reachable states are found when the control is required to satisfy a norm constraint. The results are then applied to a specific example.

CHAPTER 2

PARTIAL DIFFERENTIAL EQUATIONS IN CONTROL SYSTEMS APPLICATIONS

2.1 GENERAL DISCUSSION OF THE SYSTEMS IN THIS REPORT

This report will be concerned with the control of systems described by linear partial differential equations. Systems which occur in control systems applications can usually be described in one of the two following ways:

$$\frac{\partial y}{\partial t} = Ay + f \tag{2.1}$$

or

$$\frac{\partial^2 y}{\partial t^2} + \alpha \frac{\partial y}{\partial t} = Ay + f \tag{2.2}$$

$y(x,t)$ is a real valued function of the variables (x,t) where $x \varepsilon \Omega$ and $t \varepsilon T$

Ω = a bounded open set in E_n

E_n = n dimensional Enclidean space

T = $(0, t_1)$, a time interval with 0 arbitrarily chosen as the initial time and t_1 a final time

α = a constant

A = a spatial operator defined on some domain, $D(A)$, dense in $L_2(\Omega)$

$L_2(\Omega)$ = space of square integrable functions defined on Ω.

$D(A)$ = domain of A

$f \varepsilon L_2(\Omega x T)$

In addition to the above, a set of boundary conditions is given which can be expressed in the form

$$Uy = w$$

$$w = \text{a given function on } (\partial\Omega x T)$$

$$\partial\Omega = \text{boundary of } \Omega.$$

For example

$$A = \frac{\partial^2}{\partial x^2}$$

$$\Omega = (0,1)$$

$$\partial\Omega = \{0,1\}$$

$$Uy\ (0,t) = 0$$

$$Uy\ (1,t) = u(t), \quad u(t)\ \varepsilon\ L_2(T)$$

$$w(0,t) = 0$$

$$w(1,t) = u(t)$$

represents the partial differential equation

$$\frac{\partial y(x,t)}{\partial t} = \frac{\partial^2 y(x,t)}{\partial x^2} \qquad 0<x<1, \quad 0<t<t_1$$

$$y(0,t) = 0$$

$$y(1,t) = u(t)$$

The initial conditions of the state of the system are also assumed given. The state of the system is assumed to be as follows.

$y(x,t)$ = state of the system for equation (2.1)

$\begin{pmatrix} y(x,t) \\ \frac{\partial y(x,t)}{\partial t} \end{pmatrix}$ = state of the system for equation (2.2)

The system given in equation (2.2) can be changed to the form of that given in equation (2.1) by introducing the two component vector $z(x,t)$,

$$z(x,t) = \begin{bmatrix} y(x,t) \\ \frac{\partial y(x,t)}{\partial t} \end{bmatrix}$$

Then $z(x,t)$ satisfies the differential equation

$$\frac{\partial z}{\partial t} = A'z + f'$$

where

$$A' = \begin{bmatrix} 0 & 1 \\ A & -\alpha \end{bmatrix}.$$

$$f' = \begin{bmatrix} 0 \\ f \end{bmatrix}$$

Thus the partial differential equations to be treated can be written very briefly in the form

$$\frac{\partial y}{\partial t} = Ay + f \tag{2.3}$$

$$Uy = w$$

$$y(x,0) = y_o(x) \tag{2.4}$$

Equation (2.3) represents the partial differential equation and equation (2.4) the boundary conditions and the initial conditions.

The form of equations (2.3) and (2.4) is typical for control systems applications. The functions f and w are called the controls and it is their selection in making the state behave in some desired manner which makes the problem one in control systems. The function f is called a distributed control and the function w is called a boundary control.

It will be assumed that the system described by equations (2.3) and (2.4) is well posed. By this it is meant that (1) a solution exists, (2) it is unique, and (3) the solution depends continuously on the initial data and the control. The meaning of (3) is made more precise by the introduction of sets of norms on appropriate spaces. The function spaces to be considered are the $L_2(\Omega)$, $L_2(\Omega x T)$, and $L_2(T)$ spaces. The inner product on each of these spaces is denoted as follows.

For p, q ε $L_2(\Omega)$,

$$[p, q]_\Omega = \int_\Omega p(x)q(x)dx.$$

For f, g ε L_2 $(\Omega \times T)$

$$[f, g]_{\Omega \times T} = \int_0^{t_1} \int_\Omega f(x,t)g(x,t)dxdt.$$

For u, v ε $L_2(T)$

$$[u,v]_T = \int_0^{t_1} v(t)u(t)dt.$$

From each of these inner products, a norm follows naturally. For example,

for p ε $L_2(\Omega)$,

$$||p||_\Omega = [p, p]_\Omega^{1/2}$$

Thus, by continuous dependence on the initial data it is meant that if $y_1(x,t)$ and $y_2(x,t)$ are the response to initial conditions $y_1(x,0)$ and $y_2(x,0)$ respectively, then y_1 is close to y_2 in the $\Omega \times T$ norm provided $y_1(x,0)$ is close to $y_2(x,0)$ in the Ω norm. I.e., given $\varepsilon>0$, there is a $\delta>0$ such that

$$||y_1(x,0) - y_2(x,0)||_\Omega < \delta$$

implies

$$||y_1 - y_2||_{\Omega \times T} < \varepsilon.$$

Similar continuity in terms of the norms is implied with respect to the control terms.

The existence and uniqueness requirements of the well posed assumption of equations (2.3) and (2.4), along with the linearity, implies that if the initial condition $y_o \varepsilon L_2(\Omega)$, the distributed control $f \varepsilon L_2(\Omega \times T)$, and the boundary conditions include a control, say $u \varepsilon L_2(T)$, then the solution can be written

$$y = L_I y_o + L_\Omega f + L_B u \tag{2.5}$$

where L_I, L_Ω and L_B are linear operators.

L_I = operator from set of initial conditions to solutions

L_Ω = operator from distributed controls to solutions

and

L_B = operator from boundary controls to solutions

The operators L_I, L_Ω, and L_B are integral operators given by their corresponding Green's functions G_I, G, and G_B.

$$L_I y_o(x,t) = \int_\Omega G_I(x,t;\ \xi)\ y_o(\xi) d\xi \tag{2.6}$$

$$L_\Omega f(x,t) = \int_0^{t_1} \int_\Omega G(x,t;\ \xi,\tau)\ f(\xi,\tau) d\xi dt \tag{2.7}$$

$$L_B u(x,t) = \int_0^{t_1} G_B(x,t;\ \tau)\ u(\tau) d\tau \tag{2.8}$$

where for almost all $(x,t) \ \varepsilon\ \Omega \times T$, $G_I(x,t;\ \xi)\ \varepsilon L_2(\Omega)$ as a function of ξ, $G(x,t;\ \xi,\tau)\ \varepsilon L_2(\Omega x T)$ as a function of (ξ,τ) and $G_B(x,t;\ \tau) \varepsilon L_2(T)$ as a function of τ.

One of the important topics in the study of partial differential equations is to determine when the problem presented in equations (2.3) and (2.4) has continuous inverses L_I, L_Ω, and L_B.

The answer to this question is not easily resolved in the very general setting in which the equations have been written here. The specification of conditions under which a continuous inverse does exist has required a very abstract mathematical treatment and is generally beyond the level of rigor which engineers usually employ when attempting to solve specific problems. The approach to be taken in this dissertation is to apply the techniques used by Friedman [12] to control systems problems. That is, it is possible to discuss many properties of differential equations in general terms of linear operators. The general theory of linear operators belongs to the field of functional analysis which is a highly developed field of mathematics, and it is not the purpose of this dissertation to contribute to the theory of functional analysis or partial differential equations.

The next section will be concerned with finding the functions G_I, G, and G_B in special cases of interest in control systems applications.

2.2 HOMOGENEOUS BOUNDARY VALUE PROBLEMS

One of the most important methods for solving problems involving partial differential equations is the use of eigenfunction expansions. Its use will now be outlined for the case where

$$\frac{\partial y(x,t)}{\partial t} = Ay(x,t) + f \tag{2.9}$$

$$Uy = 0$$

$$y(x,0) = y_o(x)$$

First, the eigenvalue-eigenfunction problem associated with the above is solved. That is, find the set $\{\psi_n\}$ such that

$$A\psi_n(x) = \lambda_n \psi_n(x) \tag{2.10}$$

$$U\psi_n = 0$$

where the ψ_n's are functions of $x \varepsilon \Omega$ only.

The adjoint operator to A and the boundary conditions on which it acts are found next.

Let

A^* = adjoint operator of A

U^* = adjoint boundary condition operator

Then U^* and A^* are defined to satisfy the relation

$$[q, Ap]_\Omega = [A^*q, p]_\Omega$$

for all p such that

$$Up = 0$$

and all q such that

$$U^*q = 0.$$

The adjoint set of eigenvalues and eigenfunctions are also found from

$$A^*\phi_n = \gamma_n\phi_n \tag{2.11}$$

Again the ϕ_n's are functions of $x \in \Omega$ only. Two important relationships between the ψ_n's and ϕ_n's are

$$[\phi_n, \psi_m]_\Omega = 0 \qquad \text{if } \gamma_n \neq \lambda_m$$

and if λ_n is an eigenvalue of A, it is also an eigenvalue of A* [12, p. 199].

Since the magnitudes of the eigenfunctions are arbitrary, they may be normalized to satisfy

$$[\phi_n, \psi_m]_\Omega = \delta_{mn} \tag{2.12}$$

where

$$\delta_{mn} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

Assuming the $\{\psi_n\}$ span $L_2(\Omega)$, an arbitrary function $y \varepsilon L_2(\Omega \times T)$ can be expanded as

$$y(x,t) = \sum_{n=1}^{\infty} y_n(t)\psi_n(x) \tag{2.13}$$

where

$$y_n(t) = \int_{\Omega} y(x,t)\phi_n(x)dx \tag{2.14}$$

since, for almost all $t \varepsilon T$, $y(x,t) \varepsilon L_2(\Omega)$.

Also, expand $f(x,t)$ in terms of $\{\psi_n\}$.

$$f(x,t) = \sum_{n=1}^{\infty} f_n(t)\psi_n(x) \tag{2.15}$$

where

$$f_n(t) = \int_{\Omega} f(x,t)\phi_n(x)dx \tag{2.16}$$

Substitute the above into equation (2.9)

$$\sum_{n=1}^{\infty} \dot{y}_n(t)\psi_n(x) = \sum_{n=1}^{\infty} y_n(t)A\psi_n(x) + \sum_{n=1}^{\infty} f_n(t)\psi_n(x) \tag{2.17}$$

Using equation (2.10),

$$\sum_{n=1}^{\infty} \dot{y}_n(t)\psi_n(x) = \sum_{n=1}^{\infty} y_n(t)\lambda_n\psi_n(x) + \sum_{n=1}^{\infty} f_n(t)\psi_n(x) \tag{2.18}$$

Multiplying through equation (2.18) by $\phi_n(x)$ and integrating over Ω, the following set of ordinary differential equations results.

$$\dot{y}_n(t) = \lambda_n y_n(t) + f_n(t) \qquad n = 1, 2, \ldots \tag{2.19}$$

with initial conditions

$$y_n(0) = \int_\Omega y(x,0)\phi_n(x)dx$$

Thus the solution to this countably infinite set of ordinary differential equations, together with the expansion, equation (2.13), provide the solution to the partial differential equation, equation (2.9). Notice that the solution given by equation (2.13) satisfies the boundary conditions by the choice of the eigenfunctions ψ_n. I.e.,

$$Uy = \sum y_n(t)\, U\psi_n = 0 \tag{2.20}$$

because of equation (2.10).

This form of the solution is very valuable in engineering applications. In general, the eigenfunctions cannot be found analytically; however, a sufficient finite number to adequately approximate the infinite expansion can usually be found either numerically from equation (2.10) or experimentally from the physical model itself.

The solution to equation (2.19) is given by:

$$y_n(t) = y_n(0)e^{\lambda_n t} + \int_0^t e^{\lambda_n(t-\tau)} f_n(\tau)d\tau \tag{2.21}$$

Multiplying through equation (2.21) by $\psi_n(x)$ and summing

$$\sum_{n=1}^{\infty} y_n(t)\psi_n(x) = \sum_{n=1}^{\infty} y_n(0)e^{\lambda_n t}\psi_n(x) + \int_0^t \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} \psi_n(x) f_n(\tau)d\tau \tag{2.22}$$

Substituting equation (2.13) and equation (2.16) into equation (2.22),

$$y(x,t) = \sum_{n=1}^{\infty} y_n(0)e^{\lambda_n t}\psi_n(x) + \int_0^t \int_\Omega \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} \psi_n(x)\phi_n(\xi)f(\xi,\tau)d\xi d\tau \tag{2.23}$$

define

$$y_o(x,t) = \sum_{n=1}^{\infty} y_n(0)e^{\lambda_n t}\psi_n(x) \tag{2.24}$$

$$H(t) = \begin{cases} 0 & t<0 \\ 1 & 0<t \end{cases} \tag{2.25}$$

the value of H at t=0 is left undefined.

$$G(x,t;\ \xi,\tau) = H(t-\tau) \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} \psi_n(x)\phi_n(\xi) \tag{2.26}$$

then

$$y(x,t) = y_o(x,t) + \int_0^{t_1}\int_\Omega G(x,t;\ \xi,\tau)f(\xi,\tau)d\xi d\tau \tag{2.27}$$

$G(x,t;\ \xi,\tau)$ is the Green's function for this problem, and $y_o(x,t)$ is the response due to the initial conditions. $H(t)$ is the Heaviside unit step function.

2.3 NON-HOMOGENEOUS BOUNDARY VALUE PROBLEMS

In many engineering applications the control does not appear as shown in equation (2.9), but appears at the boundary in the form

$$\begin{aligned} \frac{\partial y(x,t)}{\partial t} &= Ay(x,t) \\ Uy &= w \end{aligned} \tag{2.28}$$

It is still useful to be able to have an eigenfunction expansion of $y(x,t)$; however, finding eigenfunctions of A with non-homogeneous boundary conditions is no longer a meaningful problem. The reason is that when A is defined to act on some domain where $Uy = 0$, A is a linear operator on that domain, e.g., if

$$Uy_1 = Uy_2 = 0$$

then

$$U(y_1 + y_2) = 0$$

and

$$A(y_1 + y_2) = Ay_1 + Ay_2$$

But if

$$\begin{aligned} Uy_1 &= w \\ Uy_2 &= w \\ U(y_1 + y_2) &= 2w \end{aligned}$$

so that the set of y on which A acts is not a linear space. There is a formal procedure which is very advantageous for transforming a non-homogeneous boundary value problems to an equivalent homogeneous b.v.p. In making the transformation, it is convenient to use generalized functions, in particular, δ functions and their derivatives. Some of the properties needed will now be given.

First, two representations for the δ functions will be given.

$$\delta(t) = H'(t) \qquad t \varepsilon T \tag{2.29}$$

and

$$\delta(x-\xi) = \sum_{n=1}^{\infty} \phi_n(\xi)\psi_n(x) \qquad x,\xi \varepsilon \Omega \tag{2.30}$$

The first is quite common. To prove the validity of the second, let $p(x)$ be an arbitrary function on Ω. Then

$$\int_\Omega p(\xi) \sum_{n=1}^{\infty} \phi_n(\xi)\psi_n(x)d\xi = \sum_{n=1}^{\infty} p_n\psi_n(x) \tag{2.31}$$

$$= p(x) \tag{2.32}$$

Now some properties of the Green's function will be shown.

$$G(x,t;\ \xi,\tau) = H(t-\tau) \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} \psi_n(x)\phi_n(\xi) \tag{2.33}$$

$$\frac{\partial G}{\partial t} = H'(t-\tau) \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} \psi_n(x)\phi_n(\xi)$$

$$+ H(t-\tau) \sum_{n=1}^{\infty} \lambda_n e^{\lambda_n(t-\tau)} \psi_n(x)\phi_n(\xi) \tag{2.34}$$

$$AG = H(t-\tau) \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} A\psi_n(x)\phi_n(\xi) \tag{2.35}$$

$$= H(t-\tau) \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} \lambda_n\psi_n(x)\phi_n(\xi) \tag{2.36}$$

therefore

$$\frac{\partial G}{\partial t} - AG = H'(t-\tau) \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} \psi_n(x)\phi_n(\xi) \tag{2.37}$$

however

$$H'(t-\tau) = 0 \text{ for } t \neq \tau \tag{2.38}$$

thus

$$\frac{\partial G}{\partial t} - AG = H'(t-\tau) \sum_{n=1}^{\infty} \psi_n(x)\phi_n(\xi) \tag{2.39}$$

or

$$DG = \delta(t-\tau)\delta(x-\xi) \tag{2.40}$$

where

$$D = \frac{\partial}{\partial t} - A \tag{2.41}$$

Next, D* and its set of boundary conditions are found.

D* must satisfy

$$\int_0^{t_1}\int_\Omega f(x,t)Dg(x,t)dxdt = \int_0^{t_1}\int_\Omega D^*f(x,t)g(x,t)dxdt \tag{2.42}$$

for all g such that

$$Ug = 0$$

$$g(x,0) = 0$$

Writing out the left side of equation (2.42), and integrating by parts,

$$\int_0^{t_1} \int_\Omega f(x,t) \left[\frac{\partial g(x,t)}{\partial t} - Ag(x,t)\right] dxdt$$

$$= \int_0^{t_1} \int_\Omega - g(x,t) \frac{\partial f(x,t)}{\partial t} dxdt + \int_\Omega [g(x,t)f(x,t)]_{t=0}^{t=t_1} dx$$

$$- \int_0^{t_1} \int_\Omega f(x,t)Ag(x,t)dxdt \tag{2.43}$$

$$= \int_0^{t_1} \int_\Omega g(x,t)\left[- \frac{\partial f(x,t)}{\partial t} - A^*f(x,t)\right]dxdt \tag{2.44}$$

thus

$$D^*f = - \frac{\partial f}{\partial t} - A^*f \tag{2.45}$$

with f such that

$$f(x,t_1) = 0$$

$$U^*f = 0 \tag{2.46}$$

Next it will be shown that D^*G has a representation as given in equation (2.40) also. In this case D^* acts with respect to the (ξ,τ) variables of G.

Take $-\frac{\partial}{\partial\tau}$ in equation (2.33),

$$-\frac{\partial}{\partial\tau}G(x,\xi;\ t,\tau) = H'(t-\tau)\sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)}\phi_n(\xi)\psi_n(x)$$

$$+ H(t-\tau)\sum_{n=1}^{\infty}\lambda_n e^{\lambda_n(t-\tau)}\phi_n(\xi)\psi_n(x) \qquad (2.47)$$

$$A^*G = H(t-\tau)\sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} A^*\phi_n(\xi)\psi_n(x) \qquad (2.48)$$

Since the eigenvalues of A* are the same as those of A,

$$A^*\phi_n = \lambda_n\phi_n$$

Thus

$$A^*G = H(t-\tau)\sum_{n=1}^{\infty}\lambda_n e^{\lambda_n(t-\tau)}\phi_n(\xi)\psi_n(x) \qquad (2.49)$$

Combining equation (2.47) with equation (2.49),

$$-\frac{\partial G}{\partial\tau} - A^*G = H'(t-\tau)\sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)}\phi_n(\xi)\psi_n(x) \qquad (2.50)$$

or

$$D^*G = \delta(x-\xi)\delta(t-\tau) \qquad (2.51)$$

Let the boundary condition operator U act on the x variable in the expansion for G in equation (2.33),

$$UG = \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)}\phi_n(\xi)U\psi_n(x) = 0 \qquad (2.52)$$

Evaluating G in equation (2.26) at $t = 0^+$

$$G(x,0^+;\ \xi,\tau) = 0 \qquad \tau\varepsilon T$$

By this it is meant that, for all $\tau\varepsilon T$,

$$\lim_{t\to 0^+} G(x,t;\ \xi,\tau) = 0$$

Similarly, U^* acting on the ξ variable yields,

$$U^*G = \sum_{n=1}^{\infty} e^{\lambda_n(t-\tau)} U^*\phi_n(\xi)\psi_n(x) = 0 \tag{2.53}$$

Evaluating G at $\tau = t_1^-$,

$$G(x,t;\ \xi,t_1^-) = 0, \qquad t\varepsilon T$$

Summarizing these results on G, for D and U acting on the (x,t) variables,

$$DG = \delta(x-\xi)\delta(t-\tau) \tag{2.54}$$

$$UG = 0$$

$$G(x,0^+;\ \xi,\tau) = 0 \tag{2.55}$$

For D^* and U^* acting on the (ξ,τ) variables

$$D^*G = \delta(x-\xi)\delta(t-\tau) \tag{2.56}$$

$$U^*G = 0$$

$$G(x,t;\ \xi,t_1^-) = 0 \tag{2.57}$$

These results will now be applied to solving the non-homogeneous b.v.p. by making use of the solution to the homogeneous b.v.p. Let the solution to

$$\begin{aligned} Dy &= g \\ Uy &= 0 \\ y(x,0) &= 0 \end{aligned} \tag{2.58}$$

be given by

$$y(x,t) = \int_0^{t_1}\int_\Omega G(x,t;\ \xi,\tau)g(\xi,\tau)d\xi d\tau \tag{2.59}$$

or

$$y = L_\Omega g \tag{2.60}$$

and suppose the problem is to solve

$$\begin{aligned} Dy &= 0 \\ Uy &= w \\ y(x,0) &= 0 \end{aligned} \tag{2.61}$$

One method of making use of the solution of equation (2.58) to solve equation (2.61) is to find a function, h, defined on $(\Omega \cup \partial\Omega) X T$ such that

$$\begin{aligned} Uh &= w \\ h(x,0) &= 0 \end{aligned} \tag{2.62}$$

Let z be the solution to

$$Dz = -(Dh) \tag{2.63}$$

$$Uz = 0$$

$$z(x,0) = 0$$

let

$$y = z + h \tag{2.64}$$

then

$$Dy = Dz + Dh \tag{2.65}$$

$$Uy = Uz + Uh$$

$$y(x,0) = z(x,0) + h(x,0)$$

thus

$$Dy = 0 \tag{2.66}$$

$$Uy = w$$

$$y(x,0) = 0$$

hence y is the solution to equation (2.61).

Suppose there exists a generalized function F on ΩxT, depending on w, satisfying

$$[f, F]_{\Omega xT} = - [f, Dh]_{\Omega xT} + [D^*f, h]_{\Omega xT} \tag{2.67}$$

for all f such that

$$U^*f = 0$$

$$f(x,t_1) = 0$$

and h such that

$$Uh = w$$

$$h(x,0) = 0$$

The first term on the right hand side of equation (2.67) is the solution to equation (2.63) if, for fixed (x,t),

$$f(\xi,\tau) = G(x,t;\ \xi,\tau) \tag{2.68}$$

i.e.,

$$z(x,t) = \int_0^{t_1} \int_\Omega G(x,t;\ \xi,\tau)\ [Dh(\xi,\tau)]d\xi d\tau \tag{2.69}$$

However, because of equation (2.57), equation (2.67) and equation (2.69) imply

$$z(x,t) = \int_0^{t_1} \int_\Omega G(x,t;\ \xi,\tau)F(\xi,\tau)d\xi d\tau$$

$$-\int_0^{t_1} \int_\Omega [D^*G(x,t;\ \xi,\tau)]h(\xi,\tau)d\xi d\tau \tag{2.70}$$

Using equation (2.56),

$$z(x,t) = \int_0^{t_1}\int_\Omega G(x,t;\ \xi,\tau)F(\xi,\tau)d\xi d\tau - h(x,t) \tag{2.71}$$

Since the solution to equation (2.61) is given in equation (2.64),

$$y(x,t) = \int_0^{t_1} \int_\Omega G(x,t;\ \xi,\tau)F(\xi,\tau)d\xi d\tau \tag{2.72}$$

In summary, the solution to the non-homogeneous b.v.p., equation (2.61), is the solution to the homogeneous b.v.p. with forcing term F,

$$Dy = F$$

$$Uy = 0$$

$$y(x,0) = 0$$

where F is a generalized function, depending on w, satisfying equation (2.67).

2.4 EXAMPLE

A simple example will demonstrate the use of this method. Consider the heat equation given by

$$\frac{\partial y(x,t)}{\partial t} = \frac{\partial^2 y(x,t)}{\partial x^2} \tag{2.73}$$

$$y(x,0) = 0$$

$$y(0,t) = 0$$

$$y(1,t) = u(t)$$

In this case,

$$\Omega = (0,1)$$

$$T = (0,t_1)$$

$$A = \frac{d^2}{dx^2}$$

$$u(t) = \text{control}$$

First find the eigenvalues and eigenfunctions for the homogeneous boundary value problem,

$$\frac{d^2\psi(x)}{dx^2} = \lambda\psi(x) \tag{2.74}$$

$$\psi(0) = \psi(1) = 0$$

The general solution is

$$\psi(x) = a \sin \sqrt{-\lambda}\ x + b \cos \sqrt{-\lambda}\ x$$

The non-zero solutions satisfying the boundary conditions occur when

$$\lambda = -n^2\pi^2 \qquad n = 1, 2, \ldots$$

$$b = 0$$

In order to make the eigenfunctions have unit magnitude on (0,1), let

$$a = \sqrt{2}$$

Therefore the eigenfunctions are

$$\psi_n(x) = \sqrt{2}\ \sin n\pi x \tag{2.75}$$

This set forms an orthonormal complete set on (0,1).

Next the adjoint operator A* and its set of boundary conditions are found to satisfy,

$$\int_0^1 p(x)Aq(x)dx = \int_0^1 A^*p(x)q(x)dx \tag{2.76}$$

for all q(x) such that

$$q(0) = q(1) = 0$$

Writing out the left hand sie of equation (2.76) and integrating by parts

$$\int_0^1 p(x)\,\frac{d^2q(x)}{dx^2}\,dx = p(x)\,\frac{dq(x)}{dx}\Big|_0^1 - q(x)\,\frac{dp(x)}{dx}\Big|_0^1 + \int_0^1 q(x)\,\frac{d^2p(x)}{dx^2}\,dx \qquad (2.77)$$

Therefore equation (2.76) is satisfied by

$$A^*p(x) = \frac{d^2p(x)}{dx^2}$$

with the boundary conditions

$$p(0) = p(1) = 0$$

Next, a generalized function, F, is found satisfying equation (2.67) for all f such that

$$f(0,t) = f(1,t) = f(x,t_1) = 0$$

and all h such that

$$h(0,t) = 0$$

$$h(1,t) = u(t)$$

$$h(x,0) = 0$$

Writing out the second term on the right hand side of equation (2.67),

$$[D^*f,h]_{\Omega xT} = \int_0^{t_1}\int_0^1 [-\frac{\partial f(\xi,\tau)}{\partial\tau} - \frac{\partial^2 f(\xi,\tau)}{\partial\xi^2}]h(\xi,\tau)d\xi d\tau$$

Integrating by parts,

$$[D^*f,h]_{\Omega xT} = \int_0^1 [-f(\xi,\tau)h(\xi,\tau)]_{\tau=0}^{\tau=t_1} d\xi$$

$$-\int_0^{t_1} [h(\xi,\tau)\frac{\partial f(\xi,\tau)}{\partial\xi} - f(\xi,\tau)\frac{\partial h(\xi,\tau)}{\partial\xi}]_{\xi=0}^{\xi=1} d\tau$$

$$+\int_0^{t_1}\int_0^1 f(\xi,\tau)[\frac{\partial h(\xi,\tau)}{\partial\tau} - \frac{\partial^2 h(\xi,\tau)}{\partial\xi^2}]d\xi d\tau \tag{2.78}$$

Using the fact that the final term on the right hand side of equation (2.78) is $[f, Dh]_{\Omega xT}$, and substituting the boundary conditions on f and g in the remaining terms,

$$-[f, Dh]_{\Omega xT} + [D^*f, h] = -\int_0^{t_1} u(\tau)\frac{\partial f(1,\tau)}{\partial\xi} d\tau \tag{2.79}$$

If

$$F(x,t) = u(t)\delta'(x-1) \tag{2.80}$$

then

$$[F,f]_{\Omega xT} = \int_0^{t_1}\int_0^1 u(\tau)\delta'(\xi-1)f(\xi,\tau)d\xi d\tau$$

Integrating by parts,

$$[F,f]_{\Omega xT} = \int_0^{t_1} u(\tau)\left[f(\xi,\tau)\delta(\xi-1)\Big|_{\xi=0}^{\xi=1} - \int_0^1 \frac{\partial f(\xi,\tau)}{\partial\xi}\delta(\xi-1)d\xi\right] d\tau$$

$$= -\int_0^{t_1} u(\tau)\frac{\partial f(1,\tau)}{\partial\xi} d\tau \tag{2.81}$$

Since the right hand side of equation (2.81) is the same as equation (2.79), the generalized function, F, in equation (2.80), satisfies equation (2.67) for all appropriate f and h. Thus, the equivalent problem to that given in equation (2.73) is

$$\frac{\partial y(x,t)}{\partial \tau} = \frac{\partial^2 y(x,t)}{\partial x^2} + u(t)\delta'(x-1) \tag{2.82}$$

$$y(x,0) = 0$$

$$y(0,t) = 0$$

$$y(1,t) = 0$$

Expand y(x,t) in terms of the eigenfunctions of the homogeneous equation,

$$y(x,t) = \sum_{n=1}^{\infty} y_n(t)\psi_n(x) \tag{2.83}$$

Substitute equation (2.83) into equation (2.82),

$$\sum_{n=1}^{\infty} \dot{y}_n(t)\psi_n(x) = \sum_{n=1}^{\infty} y_n(t)\psi_n''(x) + u(t)\delta'(x-1) \tag{2.84}$$

From equation (2.75),

$$\psi_n''(x) = -n^2\pi^2\psi_n(x)$$

Substitute the above into equation (2.84)

$$\sum_{n=1}^{\infty} \dot{y}_n(t)\psi_n(x) = -\sum_{n=1}^{\infty} n^2\pi^2 y_n(t)\psi_n(x) + u(t)\delta'(x-1) \tag{2.85}$$

Multiply through equation (2.85) by $\psi_m(x)$ and integrate over the interval (0,1).

Using the properties,

$$\int_0^1 \psi_m(x)\psi_n(x)dx = \delta_{mn}$$

and

$$\int_0^1 \psi_m(x)\delta'(x-1)dx = -\psi_m'(1) = -(-1)^m\sqrt{2}\, m\pi$$

The following infinite set of ordinary differential equations results.

$$\dot{y}_m(t) = -m^2\pi^2 y_m(t) - (-1)^m m\pi\sqrt{2}\, u(t) \tag{2.86}$$

$$m = 1, 2, \ldots$$

The initial conditions are

$$y_m(0) = 0$$

This is a useful form for control systems problems. It appears in the usual form of linear control systems problems in finite dimensional control systems. Since there are infinitely many equations to be satisfied in this case, the appropriate term infinite dimensional control system is applied.

The solution to equation (2.86) is given by

$$y_m(t) = -(-1)^m m\pi\sqrt{2}\int_0^t e^{-m^2\pi^2(t-\tau)}u(\tau)d\tau \tag{2.87}$$

$$m = 1, 2, \ldots$$

Thus the solution to equation (2.73) is given by the expansion in equation (2.83) with $y_n(t)$ given in equation (2.87).

The heat equation, equation (2.73), also furnishes a good example of the possibility of ill posed problems arising in partial differential equations. First, look at the solution to the initial value problem with homogeneous boundary conditions. Let $y(x,t)$ be the solution to equation (2.73) with

$$y(x,0) = y_o(x)$$

$$y(0,t) = y(1,t) = 0$$

Then the solution, given in equation (2.23), is

$$y(x,t) = \sum_{n=1}^{\infty} y_n(0)\ e^{-n^2\pi^2 t}\ \sqrt{2}\ \sin n\pi x$$

where

$$y_n(0) = \int_0^1 y_o(x)\ \sqrt{2}\ \sin n\pi x dx$$

Let $y_1(x,t)$ and $y_2(x,t)$ be two states resulting from arbitrary initial states $y_1(x,0)$ and $y_2(x,0)$ and compute the norm of their difference at time $t = t_1$.

$$||y_1(x,t_1) - y_2(x,t_2)||_\Omega^2 = \int_0^1 [y_1(x,t_1) - y_2(x,t_1)]^2\, dx$$

$$= \int_0^1 \left\{ \sum_{n=1}^{\infty} [y_{n1}(0)-y_{n2}(0)] e^{-n^2\pi^2 t_1} \sqrt{2} \sin n\pi x \right\}^2 dx$$

$$= \int_0^1 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} [y_{n1}(0)-y_{n2}(0)][y_{m1}(0)-y_{m2}(0)]$$

$$e^{-(m^2 + n^2)\pi^2 t_1} 2 \sin n\pi x \sin m\pi x \quad dx$$

$$= \sum_{m=1}^{\infty} [y_{n1}(0) - y_{n2}(0)]^2 e^{-2m^2\pi^2 t_1}$$

$$\leq \sum_{m=1}^{\infty} [y_{n1}(0) - y_{n2}(0)]^2$$

$$= ||y_1(x,0) - y_2(x,0)||_\Omega^2$$

Therefore the requirement of the continuity of state at time t_1 due to the initial conditions is seen to hold. Thus $y_1(x,0)$ close to $y_2(x,0)$ at $t=0$ implies that at future times t_1, the resulting states $y_1(x,t_1)$ and $y_2(x,t_1)$ will be close also. Now suppose the reverse problem posed. Given the state at time $t_1>0$, is it possible to determine what the state of the system was at time $t=0$? The answer is that it is only possible to give the initial state if the state at time t_1 is known exactly. If two states are close at time t_1, it says nothing about how close the states were at time $t=0$. For example, let two states at time $t=0$ be given by

$$y_1(x,0) = \sin \pi x$$

$$y_2(x,0) = \sin \pi x + C \sin N\pi x$$

The states at time t_1 are

$$y_1(x,t_1) = e^{-\pi^2 t_1} \sin \pi x$$

$$y_2(x,t_1) = e^{-\pi^2 t_1} \sin \pi x + Ce^{-N^2\pi^2 t_1} \sin N\pi x$$

Therefore

$$||y_1(x,t_1) - y_2(x,t_1)||_\Omega = |C|e^{-N^2\pi^2 t_1}$$

Thus for $t_1>0$, and arbitary C, it is possible to make $y_2(x,t_1)$ arbitarily close to $y_1(x,t_1)$ by picking N sufficiently large. However at $t=0$,

$$||y_1(x,0) - y_2(x,0)|| = |C|$$

Thus it is possible for two states at time $t=0$ to be arbitrarily far apart by making $|C|$ arbitarily large even though at time t_1 the states may be made arbitarily close by choosing N sufficiently large.

PRECEDING PAGE BLANK NOT FILMED.

CHAPTER 3

CONTROLLABILITY

3.1 INTRODUCTION

This chapter will be concerned with some of the abstract ideas of controllability, and also a topic closely related to controllability which has been termed minimum effort control systems or minimum energy control systems. These systems will be treated here by the introduction of the pseudo-inverse of an operator which is a generalization of the pseudo-inverse of a matrix.

3.2 CONTROLLABILITY OF DISTRIBUTED PARAMETER SYSTEMS

For the following, consider the partial differential equation given by equations (2.3) and (2.4), and assume $w=0$ and $y_o(x) = 0$. The solution is then given by equation (2.5),

$$y = L_\Omega f \tag{3.1}$$

L_Ω is an inverse operator to the differential operator $\frac{\partial}{\partial t} - A$ in the sense that

$$\left(\frac{\partial}{\partial t} - A\right) L_\Omega f = f \tag{3.2}$$

for all $\quad f \varepsilon L_2(\Omega x T)$

and

$$L_\Omega\left(\frac{\partial}{\partial t} - A\right) y = y \tag{3.3}$$

for all $\quad y \varepsilon\ D\left(\frac{\partial}{\partial t} - A\right)$

If $y(x,t)$ is any state resulting from a control f given by equation (3.1), equation (3.2) implies that this state y must be in the domain of the operator $\frac{\partial}{\partial t} - A$. Thus at any time $t \varepsilon T$, the state $y(x,t)$ must be in the domain of the operator A. Since A is a differential operator, it can be seen that it is not possible for the set of states to be the whole space in which the states lie, i.e., $L_2(\Omega)$. The set of states which can be achieved is at most a dense subset, D(A), of the whole space. This is quite different from the finite dimensional case where A is a matrix operator, and the set of states which can be reached is the whole state space. The generalization of the concepts of controllability for infinite dimensional systems will now be discussed.

Let the time be fixed at t_1, and the state be given as $y(x,t_1)$. Let

Y = state space

U = space of controls

L = linear operator from the control space to the state space at fixed time t_1.

The system defined by the spaces Y and U and the operator L will be called the system (L, U, Y).

For example, if the relation between the control f and the state y is given by

$$y(x,t_1) = \int_0^{t_1} \int_\Omega G(x,t_1;\ \xi,\tau) f(\xi,\tau) d\xi d\tau$$

the operator L is given by

$$Lf(x,t_1) = \int_0^{t_1} \int_\Omega G(x,t_1;\ \xi,\tau) f(\xi,\tau) d\xi d\tau$$

The control space U is

$$U = L_2(\Omega x T)$$

The state space is

$$Y = L_2(\Omega)$$

The case of boundary control can also be considered under this general discussion. If the relation between the control u and the state y is given by

$$y(x,t_1) = \int_0^{t_1} G_B(x,t;\ \tau)u(\tau)d\tau$$

The operator L is given by

$$Lu(x,t_1) = \int_0^{t_1} G_B(x,t_1;\ \tau)u(\tau)d\tau$$

The control space U is

$$U = L_2(T)$$

The state space is

$$Y = L_2(\Omega)$$

The problem of controllability is to determine if, for a given desired state $y_d \varepsilon Y$, there is a control $u \varepsilon U$ and a finite time t_1 such that

$$y_d = Lu \tag{3.4}$$

In addition to the requirement that u lie in U, there may be additional constraints, for example requiring that u satisfy a constraint in the magnitude of its norm.

The reachable set of states is natural to define as follows. Let

$$R = \{y \varepsilon Y;\ y=Lu \quad u\text{:admissible}\}$$

where u belongs to a set of controls which are to be termed admissible. The set R is called the reachable set. Two particularly important classes of admissible controls to consider arise when the control occurs at the boundary with $U = L_2(T)$ and the constraints are (i) $\int_0^{t_1} u^2(t)dt < \infty$ and (ii) $\int_0^{t_1} u^2(t)dt \leq M$ for some positive constant M. The first is just the requirement that $u \varepsilon U$; the second is a norm constraint on u. The aspects of controllability under the first constraint will be the topic of the present chapter; the reachable set under the second constraint will be the topic of the next chapter.

When the set of admissible controls is the whole space U, it can be seen that the reachable set is the range of the operator L.

Define

$$R(L) = \text{range of } L$$

$$R(L) = \{y \varepsilon Y; \; y = Lu, \; u \varepsilon U\}$$

As it was noted previously, the range of L is in general not closed.

Define

$$\overline{R(L)} = \text{closure of the range of } L$$

Now the definition of complete controllability is made as follows.

Definition 2.1. Complete Controllability. The system (L, U, Y) is completely controllable if and only if $\overline{R(L)} = Y$.

The physical interpretation of this definition is that although not every point in Y can be reached with a control from U, $R(L)$ is dense in Y. Therefore it is possible to come arbitarily close to any state in Y if the system is completely controllable. Closeness is meant in the sense of the norm in Y which is a Hilbert space.

Since the operator L is a linear operator from one Hilbert space, U, to another Hilbert space, Y, it is possible to obtain several results concerning complete controllability based on theorems which are easily obtained in functional analysis.

Let the inner products on U and Y be denoted by $[u,v]_U$ for $u,v\varepsilon U$ and $[x,y]_Y$ for $x,y\varepsilon Y$.

Define the adjoint operator as a linear mapping from Y to U.

$$L^*:Y \to U$$

such that

$$[y,Lu]_Y = [L^*y,u]_U$$

for all

$$u\varepsilon U$$

$$y\varepsilon Y$$

Define the null space of L as $N(L)$.

$$N(L) = \{u\varepsilon U; \quad Lu = 0\}$$

Define the orthogonal complement of $\overline{R(L)}$ as $\{\overline{R(L)}\}^{\perp}$

$$\overline{\{R(L)\}}^{\perp} = \{y \varepsilon Y;\ [y,Lu]_Y = 0 \text{ for all } u \varepsilon U\}$$

Since Y is a Hilbert space, it has the following direct sum decomposition [13, p. 246].

$$Y = \overline{\{R(L)\}} \oplus \overline{\{R(L)\}}^{\perp} \tag{3.5}$$

By this it is meant that for each $y \varepsilon Y$, there are unique elements y_1 and y_2 such that $y_1 \varepsilon \overline{R(L)}$, $y_2 \varepsilon \overline{\{R(L)\}}^{\perp}$ and

$$y = y_1 + y_2$$

From the above definitions, the following relationships are easily proven to be true [13, p. 250].

$$\overline{R(L)} = N(L^*)^{\perp} \tag{3.6}$$

$$\overline{\{R(L)\}}^{\perp} = N(L^*) \tag{3.7}$$

$$\overline{R(L)} = \overline{R(LL^*)} \tag{3.8}$$

$$N(L^*) = N(LL^*) \tag{3.9}$$

A theorem regarding complete controllability can now be stated which is a generalization of that given by Kalman, et.al., [7], for finite dimensional systems.

Theorem 2.1 The system (L, U, Y) is completely controllable if and only if $N(LL^*) = \{0\}$. ($\{0\}$ is the set consisting of only the zero element)

Proof. By the direct sum decomposition in equation (3.5),

$$Y = \overline{\{R(L)\}} \oplus \overline{\{R(L)\}}^{\perp}$$

Therefore, $Y = \overline{R(L)}$ if and only if $\overline{\{R(L)\}}^{\perp} = \{0\}$.

By equation (3.7), $\overline{\{R(L)\}}^{\perp} = N(L^*)$. Therefore, by equation (3.9), $\overline{\{R(L)\}}^{\perp} = N(LL^*)$. Hence the system (L, U, Y) is completely controllable if and only if $N(LL^*) = \{0\}$. Q.E.D.

3.3 PSEUDO-INVERSE OF L

Some further generalizations of the results in [7] can be made. If there is a $u \varepsilon U$ satisfying equation (3.4) for a given y_d, it is usually not unique. It is of interest to determine which u of all those satisfying equation (3.4) has the minimum norm. Since the Hilbert space chosen to work with is $L_2(T)$, the square of the norm is proportional to the energy required so that the minimum norm control may also be called the minimum energy or minimum effort control. It is possible that y_d is not in the range of L. In this case the decomposition of y_d is given by

$$y_d = y_1 + y_2 \tag{3.10}$$

with

$$y_1 \varepsilon \overline{R(L)}$$

$$y_2 \varepsilon \{\overline{R(L)}\}^{\perp}$$

This implies either there is a non-zero component y_2 or y_1 is actually a limit point of R(L) and not in R(L).

Let

$$R(L)' = \text{limit points of the range of } L \text{ but not in } R(L).$$

If $y_1 \varepsilon R(L)'$, then it is possible to come arbitrarily close to y_1 with controls from U, but it is impossible to achieve y_1 exactly. In this case, seeking a u of minimum norm is no longer meaningful and is one of the difficulties which are encountered in infinite dimensional systems which does not arise in the finite dimensional problems. If $y_1 \varepsilon R(L)$, and y_2 in equation (3.10) is non-zero, the question of which $y \varepsilon R(L)$ is closest to y_d can properly be asked. It is quite easy to see that y_1 is the closest. That is,

$$||y_1 - y_d||_Y \leq ||y - y_d||_Y \tag{3.11}$$

for all $y \varepsilon R(L)$.

The proof of the inequality in (3.11) is as follows:

$$||y - y_d||_Y^2 = [y - y_d, y - y_d]_Y$$

$$= [(y - y_1) - y_2, (y - y_1) - y_2]_Y$$

$$= ||y - y_1||_Y^2 - 2[y_2, (y - y_1)]_Y + ||y_2||_Y^2$$

But

$$[y_2, (y - y_1)]_Y = 0$$

since

$$y - y_1 \varepsilon R(L)$$

and

$$y_2 \varepsilon \{\overline{R(L)}\}^{\perp}$$

Therefore

$$||y - y_d||_Y^2 = ||y - y_1||_Y^2 + ||y_2||_Y^2$$

and is minimum when $y = y_1$.

A decomposition of the space of controls, U, is also possible in terms of N(L) and $N(L)^{\perp}$

$$U = N(L) \oplus N(L)^{\perp}$$

Therefore for all $y_1 \varepsilon R(L)$, and for all $u \varepsilon U$ such that

$$y_1 = Lu \tag{3.12}$$

u has the unique representation

$$u = u_1 + u_2$$

where

$$u_1 \varepsilon N(L)^{\perp}$$

and

$$u_2 \varepsilon N(L)$$

By a similar argument to the above, the u of minimum norm satisfying equation (3.12) is u_1. The definition of a pseudo-inverse of the operator L can now be made which is an extension of the idea of the pseudo-inverse of a matrix.

Let

$$L^{+} = \text{pseudo-inverse of } L$$

The desired properties of L^{+} are that it be a linear mapping from Y to U such that for $y \varepsilon Y$, and

$$u = L^{+}y$$

u is such that $||y-Lu||$ is minimum, and if

$$y_1 = Lu \tag{3.13}$$

u is the element of U having minimum norm satisfying equation (3.13). As it has been pointed out, the range of L is not closed and therefore it is not possible to define the pseudo-inverse L^{+} on the whole space Y such that it has the properties listed above. The pseudo-inverse for L is defined as follows.

Definition 2.2. Pseudo-inverse of L

The pseudo-inverse of L is a mapping, L^{+}, from $R(L) \oplus \overline{\{R(L)\}}^{\perp}$ to U such that for

$$y_1 \varepsilon R(L)$$

$$L^{+}y_1 = u_1$$

where $u_1 \varepsilon N(L)^{\perp}$ and $Lu_1 = y_1$

and for $y_2 \varepsilon \overline{\{R(L)\}}^{\perp}$

$$L^{+}y_2 = 0$$

This definition does give the operator L^{+} the desired properties, but it is useful to be able to express it more directly in terms of L and L^*. By the definition of L^{+},

$$R(L^{+}) = N(L)^{\perp} \tag{3.14}$$

Interchanging the role of L and L^* in equation (3.6) and using the result in equation (3.14),

$$R(L^+) = \overline{R(L^*)} \tag{3.15}$$

equation (3.8) and equation (3.15) together imply

$$R(L^+) = \overline{R(L^*L)}$$

Thus the problem of finding the pseudo inverse of L, given y_d, is to find the $u \varepsilon \overline{R(L^*L)}$ such that equation (3.13) is satisfied where y_d has the decomposition given by equation (3.10). In order to find the solution, let $\{\gamma_i\}$ be the non-zero eigenvalues of L^*L and $\{\phi_i\}$ the corresponding eigenfunctions.

$$L^*L\phi_i = \gamma_i\phi_i \tag{3.16}$$

Since the γ_i are non-zero, equation (3.16) implies $\phi_i \varepsilon R(L^*L)$ for all i. Also, the ϕ_i form an orthonormal set because of the self-adjointness of L^*L. Assuming that u can be expanded in terms of the $\{\phi_i\}$ as

$$u = \sum_{i=1}^{\infty} u_i\phi_i \tag{3.17}$$

then it follows that $u \varepsilon \overline{R(L^*L)}$

The above expansion is valid for the purpose in which it is going to be used later, provided the operator L^*L is completely continuous (or compact, as it is also called) [13, p. 336]. The question of whether the operator L or L^*L is completely continuous in connection with partial differential equations is one which is currently receiving a great deal of attention by mathematicians. As stated previously, it is not the purpose of this dissertation to attempt to answer questions such as these, and the assumption is simply made that the expansion in equation (3.17) is valid for all $u \varepsilon \overline{R(L^*L)}$.

Before concluding with the solution to equation (3.13), the following lemma is needed.

Lemma 3.1 $\qquad L^*y_d = L^*y_1$

y_d is given by the expansion in equation (3.10); therefore,

$$L^*y_d = L^*y_1 + L^*y_2$$

since $y_2 \varepsilon \overline{\{R(L)\}}^{\perp}$, equation (3.7) implies

$$L^*y_2 = 0$$

hence

$$L^*y_d = L^*y_1 \tag{3.18}$$

Now, suppose there is a u satisfying equation (3.13). Using the expansion in equation (3.17),

$$L\left(\sum_{i=1}^{\infty} u_i \phi_i\right) = y_1$$

Operating on both sides of the above by L* and using equation (3.18),

$$L^*L\left(\sum_{i=1}^{\infty} u_i \phi_i\right) = L^*y_d \tag{3.19}$$

Using Equation (3.16),

$$\sum_{i=1}^{\infty} \gamma_i u_i \phi_i = L^* y_d$$

Taking the inner product of the above with ϕ_j in U and using the orthonormality of the $\{\phi_i\}$, it is possible to solve for u_j,

$$u_j = \frac{1}{\gamma_j} [\phi_j, L^* y_d]_U$$

Thus the pseudo inverse of L is

$$L^+ y_d = \sum_{j=i}^{\infty} \frac{1}{\gamma_j} [\phi_j, L^* y_d]_U \phi_j$$

3.4 COMMENTS ON RELATED WORK

A general summary of the results available on controllability for finite dimensional control systems can be found in Kalman, et. al., [7]. In that paper, the minimum energy control is found by making use of the pseudo-inverse of a matrix. For finite dimensional control systems, the operator LL^* is simply a matrix, for which the notion of a pseudo-inverse was first introduced by Penrose [14]. The relationship between the pseudo-inverse of LL^* and the pseudo-inverse of L given in definition 2.2 for finite dimensional control systems is contained in the following lemma.

Lemma 3.2. If the state space Y is finite dimensional, $L^+ = L^*(LL^*)^+$.

Proof. For finite dimensional control systems, $\overline{R(L)} = R(L)$. Thus the decomposition of Y in equation (3.5) becomes

$$Y = R(L) \oplus R(L)^{\perp}$$

Let $y_1 \varepsilon R(L)$. Then by definition 2.2,

$$L^+ y_1 = u_1 \tag{3.20}$$

where

$$u_1 \varepsilon N(L)^{\perp}$$

and

$$Lu_1 = y_1 \tag{3.21}$$

Let

$$L^*(LL^*)^+y_1 = u_o \qquad (3.22)$$

The first step is to show

$$u_o = u_1$$

Let

$$y_o = (LL^*)^+y_1 \qquad (3.23)$$

Since $y_1 \varepsilon R(L)$, by equation (3.8),

$$y_1 \varepsilon R(LL^*) \qquad (3.24)$$

Therefore, by definition (2.2),

$$y_o \varepsilon N(LL^*)^{\perp} \qquad (3.25)$$

and

$$LL^*y_o = y_1 \qquad (3.26)$$

By equations (3.22) and (3.23)

$$u_o = L^*y_o \qquad (3.27)$$

hence

$$u_o \varepsilon R(L^*)$$

Therefore by equation (3.6), interchanging the role of L and L^*,

$$u_o \varepsilon N(L) \tag{3.28}$$

Operating L onto equation (3.27),

$$Lu_o = LL^*y_o \tag{3.29}$$

Therefore, because of equation (3.26),

$$Lu_o = y_1 \tag{3.30}$$

Subtracting terms in equations (3.20) and (3.30) implies

$$L(u_o - u_1) = 0$$

Since $u_o - u_1 \varepsilon N(L)^{\perp}$,

$$u_o - u_1 = 0$$

and completes the first step of the proof for $y_1 \varepsilon R(L)$.

Next let $y_2 \varepsilon R(L)^{\perp}$. By definition 2.2,

$$L^+ y_2 = 0$$

Because of equation (3.8), $y_2 \varepsilon R(LL^*)$,

hence

$$(LL^*)^+ y_2 = 0$$

also. Thus

$$L^*(LL^*)^+ y_2 = 0$$

Since

$$L^+y_1 = L^*(LL^*)^+y_1$$

$$\text{for} \quad y_1 \varepsilon R(L)$$

and

$$L^+y_2 = L^*(LL^*)^+y_2$$

$$\text{for} \quad y_2 \varepsilon R(L)^{\perp}$$

the proof of the lemma is complete.

The pseudo-inverse of matrices has had many applications and recently the pseudo-inverse of more general linear operators have been defined and studied, e.g., Loud [15]. Loud applies the generalized inverse to differential operators which do not have a unique solution. The definition given in this dissertation is an extension to infinite dimensional spaces of that given by Zadeh and Desoer [16] for matrices.

The study of minimum effort control systems by the methods of functional analysis has been carried out by many people. The work of Hsieh [8, 17] and Balakrishnan [18] is the most closely related to that appearing here. They use the fact that the minimum effort control, u, must satisfy the equation

$$L^*Lu = L^*y_d \tag{3.31}$$

The above is equation (3.19). They use the eigenfunctions of L^*L to obtain an expansion of u and solve equation (3.31). This was the method used following equation (3.19).

The only related work in the engineering literature on the controllability of distributed parameter systems known to the author is by Wang [1] and Brogan [2]. Neither of them gives the definition of complete controllability as given in definition 2.1, and both are led to some erroneous conclusions.

Wang made the statement that the existence of an inverse of the operator L^*L was necessary and sufficient for complete controllability. Also, the statement was made in a footnote the existence of $(LL^*)^{-1}$ was an equivalent condition although no proof was given. It can easily be shown that the first statement is not true even for finite dimensional control systems. Definition 2.1 reduces to the ordinary definition for complete controllability as any reasonable definition should. The second statement is equivalent to theorem 2.1 which has been proven here. Notice that $N(LL^*) = \{0\}$ is equivalent to the statement that the inverse of LL^* exists [13, p. 18]. A simple finite dimensional counter-example of a completely controllable system will show that in general $(L^*L)^{-1}$ does not exist. Consider the scalar system described by the first order ordinary differential equation

$$\frac{dy}{dt} = \alpha y + u \qquad 0<t<t_1 \tag{3.32}$$

$$y(0) = 0$$

In this case the control space, U, is

$$U = L_2(T)$$

The state space, Y, is

$$Y = E_1$$

The linear operator, L, mapping U into Y is

$$Lu = \int_0^{t_1} e^{\alpha(t_1-\tau)} u(\tau)d\tau$$

The inner product in U is

$$[u, v]_U = \int_0^{t_1} u(t)v(t)dt, \qquad u,v\varepsilon U$$

The inner product in Y is

$$[x,y]_Y = xy \qquad x,y\varepsilon Y$$

The adjoint operator to L is found to satisfy

$$[x,Lu]_Y = [L^*x,u]_U \tag{3.33}$$

for all $x\varepsilon Y$ and all $u\varepsilon U$.

The left hand side of equation (3.33) is

$$[x,Lu]_Y = x \int_0^{t_1} e^{\alpha(t_1-\tau)} u(\tau)d\tau \tag{3.34}$$

The right hand side of equation (3.33) is

$$[L^*x, u]_U = \int_0^{t_1} u(t)L^*x \, dt \tag{3.35}$$

To obtain the equality required in equation (3.33),

$$L^*x = e^{\alpha(t_1-t)} x$$

and belongs to U as a function of t.

The operator LL* is given by

$$LL^*x = \int_0^{t_1} e^{2\alpha(t_1-\tau)} x d\tau$$

$$= \frac{-1}{2\alpha}(1 - e^{2\alpha t_1})x$$

Therefore in this simple case, the operator LL* is just a scalar,

$$LL^* = \frac{1}{2\alpha}(e^{2\alpha t_1} - 1)$$

Thus LL* has an inverse provided $\alpha \neq 0$ and $0 < t_1$ and the system is completely controllable. However the operator L*L is a mapping from U to U and is given by

$$L^*Lu = e^{\alpha(t_1-t)} \int_0^{t_1} e^{\alpha(t_1-\tau)} u(\tau) d\tau$$

L*Lu belongs to U as a function to t. In this case it is possible to find a non-zero u such that

$$L^*Lu = 0$$

For example,

$$u(t) = \begin{cases} +1 & 0 < t < t' \\ -1 & t' \leq t < t_1 \end{cases}$$

where

$$t' = -\frac{1}{\alpha} \ln \left[\frac{1+e^{-\alpha t_1}}{2}\right]$$

will do. Hence $(L^*L)^{-1}$ does not exist. This is the case in general since there is usually no unique control which will transfer the state from the origin to the desired final state. I.e., it is usually possible to find a non-zero u such that

$$Lu = 0$$

This implies

$$L^*Lu = 0$$

for non-zero u and that $(L^*L)^{-1}$ seldom exists.

Sakawa [19] studied a particular distributed parameter control system which is a slight variation of the problem in the example of Chapter 2. He found the Green's function for the system and then proceeded to derive a necessary condition which the control must satisfy in order to achieve a desired final state by use of variational methods. His resulting necessary condition was equation (3.31) after the appropriate definitions of terms used here. The unfortunate circumstance which arises in distributed parameter systems is that equation (3.31) is an integral equation which has to be solved, and in general it is very difficult to obtain an analytical result. Sakawa obtained numerical results through the use of linear programming by making a discrete approximation of the original partial differential equation and adding a magnitude constraint on the control.

3.5 SUMMARY

The new results of this chapter are on controllability of distributed parameter systems and the study of minimum energy control systems by the introduction of the pseudo-inverse of the operator L. A motivation for the definition of complete controllability in definition 2.1 has been given. This definition includes the special case of finite dimensional control systems. A generalization of the results of Kalman et.al., [7] is contained in Theorem 2.1 and a correction of the previously reported theorem by Wang [1] has been made. The generalization of the pseudo-inverse of a matrix given by Zadeh and Desoer [16] has been made to include the linear operators associated with distributed parameter systems. This method is a new approach to the study of minimum energy control systems and was shown to include the results previously presented for finite dimensional control systems by Kalman, et.al., [7] and a special case of a distributed parameter system by Sakawa [19].

CHAPTER 4

REACHABLE STATES

4.1 REACHABLE STATES WITH CONSTRAINT ON THE CONTROL

This chapter will deal with finding the set of reachable states when the magnitude of the norm of the control is constrained. It is further assumed there is only one control available and it is a function of t only. This situation arises when the control appears at the boundary as in the case of the example in Chapter 2. Also, it is assumed that the state, $y(x,t)$, is scalar valued. Although these restrictions place a limitation on the results, most distributed parameter control systems encountered in practice are included in the class just described. Thus it is assumed the system is that given by equation (2.1) with the further restriction that the control appears as follows:

$$\frac{\partial y}{\partial t} = Ay + F(x)u(t) \tag{4.1}$$

$$Uy = 0$$

$$y(x,0) = 0$$

Here, it is assumed the problem has been reduced to its equivalent homogeneous boundary value problem. Thus $F(x)$ may be a generalized function, i.e., it may include δ functions or their derivatives. The system given in equation (2.2) can also be included in this form since it can be reduced to a first order form through the introduction of a state which is a complex variable.

This will be shown in an example later. Hence, all scalar valued variables may assume complex values throughout this chapter.

As in Chapter 2, assume the state, $y(x,t)$, can be expanded in terms of the eigenfunctions of the homogeneous boundary value problem.

$$y(x,t) = \sum_{n=1}^{\infty} y_n(t)\psi_n(x)$$

where

$$A\psi_n = \lambda_n\psi_n$$

$$U\psi_n = 0 \qquad\qquad n = 1, 2, \ldots$$

Also, let the eigenvalues and eigenfunctions of the adjoint operator A^* be given by

$$A^*\phi_n = \gamma_n\phi_n$$

$$U^*\phi_n = 0 \qquad\qquad n = 1, 2, \ldots$$

The control, f, in equation (2.9) for the system described in equation (4.1) is given by

$$f(x,t) = F(x)u(t)$$

Thus the terms $f_n(t)$ given in equation (2.16) for the expansion in equation (2.15) take the form

$$f_n(t) = \int_\Omega F(x)u(t)\phi_n(x)dx$$

Let

$$b_n = \int_\Omega F(x)\phi_n(x)dx$$

Then

$$f_n(t) = b_n u(t)$$

Hence, from equation (2.21), $y_n(t)$ is found by

$$y_n(t) = \int_0^{t_1} e^{\lambda_n(t-\tau)} b_n u(\tau)d\tau$$

The object of this chapter will be to determine the reachable set of states when the control is constrained in norm. That is, it is required that

$$\left[\int_0^{t_1} |u(t)|^2 dt\right]^{\frac{1}{2}} \leq M \qquad (4.2)$$

where M is some positive constant.

Since $y(x,0) = 0$, a reachable state is meant to be reachable from the origin.

Let the time be fixed at t_1, and for brevity, let

$$C_n = y_n(t_1)$$

Since the state at time t_1 can be expanded in terms of C_n,

$$y(x,t_1) = \sum_{n=1}^{\infty} C_n \psi_n(x)$$

an equivalent definition of the state can be given. I.e., the sequence $\{C_n\}$ can be called the state of the system, and this will be done through the remainder of this chapter.

Let

$$v_n(\tau) = b_n e^{\lambda_n(t_1-\tau)}$$

The equivalent definition for the set of reachable states is the following

$$C = \left\{ \{C_n\}; \quad C_n = \int_0^{t_1} v_n(t)u(t)dt, \quad u\text{: admissible} \right\}$$

where the admissible class of controls are those satisfying the constraint in (4.2).

The problem is now to find necessary and sufficient conditions on sequences $\{C_n\}$ such that they belong to C.

The problem stated in this form is exactly the moment problem for which the solution was first given by Banach [10] and is stated in more general terms in Yosida [20]. The basic result is contained in the following theorem.

Theorem 4.1. Given (i) a sequence of functions $\{v_n\}$, where $v_n \varepsilon L_2(T)$ for each n, (ii) a sequence of complex numbers C_n, and (iii) a positive constant M, in order that there exist a function $u \varepsilon L_2(T)$ satisfying

a) $$\int_0^{t_1} |u(t)|^2 dt \le M^2$$

and

b) $$\int_0^{t_1} v_n(t)u(t)dt = C_n \tag{4.3}$$

for all n

it is necessary and sufficient that for each finite sequence of complex numbers $\{\eta_1, \ldots, \eta_N\}$, the following inequality is satisfied.

$$\left| \sum_{n=1}^{N} \eta_n C_n \right| \le M \sqrt{\int_0^{t_1} \left| \sum_{n=1}^{N} \eta_n v_n(t) \right|^2 dt} \tag{4.4}$$

The proof can be found in Banach [10] or Yosida [20, p. 106] and makes use of the Hahn-Banach theorem.

The theorem allows one to deal with a finite number of quantities in testing for the existence of a solution to an infinite number of equations in (4.3).

Some consequences of this theorem will now be given and then the results will be applied to some specific examples.

First, a more compact notation will be introduced.

Let

$$\underline{C}_N = \begin{pmatrix} C_1 \\ \vdots \\ C_N \end{pmatrix}$$

$\underline{C}_N$ is an N component column vector

Let

$$\underline{\eta}_N^* = (\bar{\eta}_1, \ldots, \bar{\eta}_N)$$

$\underline{\eta}_N^*$ is an N component row vector, where $\bar{\eta}_i$ is the complex conjugate of η_i.

$$\underline{\eta}_N^* \underline{C}_N = \sum_{n=1}^{N} \bar{\eta}_n c_n$$

and is the inner product in the complex E_N space.

The inequality in (4.2) requires that

$$|\underline{\eta}_N^* \underline{C}_N| \leq M \sqrt{\int_0^{t_1} \left| \sum_{n=1}^{N} \eta_N v_N(t) \right|^2 dt} \tag{4.5}$$

The term appearing under the radical can be simplified.

$$\int_0^{t_1} \left| \sum_{n=1}^{N} \eta_n v_n(t) \right|^2 dt$$

$$= \int_0^{t_1} \sum_{m=1}^{N} \sum_{n=1}^{N} \bar{\eta}_m \eta_n \bar{v}_m(t) v_n(t) \; dt$$

$$= \sum_{m=1}^{N} \sum_{n=1}^{N} \bar{\eta}_m \eta_n \int_0^{t_1} \bar{v}_m(t) v_n(t) dt$$

Let

$$q_{mn} = \int_0^{t_1} \bar{v}_m(t) v_n(t) dt \tag{4.6}$$

and

$$Q_N = [q_{mn}]$$

I.e., Q_N is the NXN matrix with elements q_{mn}. Thus

$$\int_0^{t_1} \left| \sum_{n=1}^{N} \eta_n v_n(t) \right|^2 dt = \sum_{m=1}^{N} \sum_{n=1}^{N} \bar{\eta}_m \eta_n q_{mn}$$

$$= \underline{\eta}_N^* \, Q_N \, \underline{\eta}_N$$

Hence Q_N is non-negative definite for all N and is positive definite if and only if $(v_1, \ldots, v_N)$ are linearly independent. Also, because of equation (4.6), Q_N is Hermitian.

The square of the magnitude of the term on the left hand side of the inequality in (4.5) can be written

$$|\underline{\eta}_N^* \underline{C}_N|^2 = \underline{\eta}_N \underline{C}_N \underline{C}_N^* \underline{\eta}_N$$

The inequality in (4.5) can be written

$$\underline{\eta}_N^* \underline{C}_N \underline{C}_N^* \underline{\eta}_N \leq M^2 \; \underline{\eta}_N^* Q_N \underline{\eta}_N$$

The dyad $C_N C_N^*$ is a matrix, so that the expression above is equivalent to

$$0 \leq \underline{\eta}_N^* \; (M^2 Q_N - \underline{C}_N \underline{C}_N^*) \; \underline{\eta}_N \tag{4.7}$$

Hence in order that the inequality in (4.4) be satisfied for all $(\eta_1, \ldots, \eta_N)$, it is necessary and sufficient that (4.7) be satisfied for all $(\eta_1, \ldots, \eta_N)$. However, this requirement is that the matrix

$M^2Q_N - \underline{C}_N\underline{C}_N^*$ be non-negative definite. This result will now be stated as a corollary to theorem 4.1.

Corollary 4.1. In order that a sequence of numbers $\{C_N\}$ belong to C, it is necessary and sufficient that the matrix

$$M^2Q_N - \underline{C}_N\underline{C}_N^* \text{ be non-negative definite for all } N.$$

The next two theorems present, in principle at least, a method to generate the set of reachable states. It will be necessary to impose the requirement that Q_N be positive definite. As stated previously, Q_N is positive definite if and only if $\{v_1, \ldots, v_N\}$ are linearly independent. If there is a linear dependence among the $\{v_1, \ldots, v_N\}$, there must be a linear dependence among the $\{C_1, \ldots, C_N\}$ also. To show that this is true, suppose there is a set of scalars $\{\alpha_1, \ldots, \alpha_N\}$ not all zero such that

$$\sum_{n=1}^{N} \alpha_n v_n = 0 \tag{4.8}$$

For any admissible u, let

$$C_n = \int_0^{t_1} v_n(t)u(t)dt, \qquad n = 1, \ldots, N.$$

Multiplying by α_n and summing implies

$$\sum_{n=1}^{N} \alpha_n C_n = \int_0^{t_1} \sum_{n=1}^{N} \alpha_n v_n(t)u(t)dt$$

But, because of equation (4.8),

$$\sum_{n=1}^{N} \alpha_n C_n = 0$$

and therefore a requirement of linear dependence on the states is imposed if the $\{v_1, \ldots, v_N\}$ are not linearly independent.

The following theorem geometrically characterizes the set of reachable states.

Theorem 4.2. If Q_N is positive definite, $M^2 Q_N - \underline{C}_N \underline{C}_N^*$ is non-negative definite if and only if $\underline{C}_N^* Q_N^{-1} \underline{C}_N \le M^2$.

Proof

(Necessity) For all $\underline{\eta}_N$,

$$0 \le M^2 \; \underline{\eta}_N^* Q_N \underline{\eta}_N - \underline{\eta}_N^* \underline{C}_N \underline{C}_N^* \underline{\eta}_N \tag{4.9}$$

Since Q_N is positive definite, Q_N^{-1} exists. Also since Q_N is Hermitian,

$$Q_N^* = Q_N$$

and

$$(Q_N^{-1})^* = Q_N^{-1}$$

Q_N^* is the complex conjugate transpose of Q_N.

Since the inequality in (4.9) must hold for all $\underline{\eta}_N$, it must hold in particular for

$$\underline{n}_N = Q_N^{-1}\underline{C}_N$$

Thus

$$0 \le M^2\ \underline{C}_N^* Q_N^{-1} Q_N Q_N^{-1}\underline{C}_N - \underline{C}_N^* Q_N^{-1}\underline{C}_N \underline{C}_N^* Q_N^{-1}\underline{C}_N$$

$$= M^2\ (\underline{C}_N^* Q_N^{-1}\underline{C}_N) - (\underline{C}_N^* Q_N^{-1}\underline{C}_N)^2$$

Since Q_N^{-1} is positive definite, for non-zero $\underline{C}_N$,

$$\underline{C}_N^* Q_N^{-1} C_N > 0$$

Therefore

$$0 \le M^2 - \underline{C}_N^* Q_N^{-1}\underline{C}_N^*$$

or

$$\underline{C}_N^* Q_N^{-1}\underline{C}_N \le M^2$$

(Sufficiency)

Since Q_N is positive definite and Hermitian, there exist positive definite Hermitian matrices $\sqrt{Q_N}$ and $\sqrt{Q_N^{-1}}$ such that

$$(\sqrt{Q_N})\ (\sqrt{Q_N}) = Q_N$$

$$(\sqrt{Q_N^{-1}})\ (\sqrt{Q_N^{-1}}) = Q_N^{-1}$$

and

$$(\sqrt{Q_N}) \quad (\sqrt{Q_N^{-1}}) = I_N$$

I_N is the identity matrix.

Starting with the identity, for arbitrary $\underline{n}_N$,

$$\underline{n}_N^* \underline{C}_N = (\sqrt{Q_N^{-1}} \sqrt{Q_N} \underline{n}_N^*) \underline{C}_N$$

it follows that

$$\underline{n}_N^* \underline{C}_N = \sqrt{Q_N} \quad \underline{n}_N^* \quad (\sqrt{Q_N^{-1}})^* \; \underline{C}_N$$

$$= \sqrt{Q_N} \quad \underline{n}_N^* \quad \sqrt{Q_N^{-1}} \; \underline{C}_N$$

Therefore

$$|\underline{n}_N^* \underline{C}_N|^2 = |\sqrt{Q_N} \quad \underline{n}_N^* \quad \sqrt{Q_N^{-1}} \underline{C}_N|^2$$

By Schwarz inequality

$$|\sqrt{Q_N} \quad \underline{n}_N^* \quad \sqrt{Q_N^{-1}} \quad \underline{C}_N|^2 \leq ||\sqrt{Q_N} \, \underline{n}_N||^2 \cdot ||\sqrt{Q_N^{-1}} \; \underline{C}_N||^2$$

Here, $||\cdot||$ indicates the Euclidean norm.

$$||\sqrt{Q_N} \quad \underline{n}_N||^2 = \underline{n}_N^* \sqrt{Q_N} \sqrt{Q_N} \quad \underline{n}_N$$

$$= \underline{n}_N^* \; Q_N \; \underline{n}_N$$

and

$$||\sqrt{Q_N^{-1}} \ \underline{C}_N||^2 = \underline{C}_N^* Q_N^{-1} \underline{C}_N$$

Thus it follows that

$$|\underline{n}_N^* \underline{C}_N|^2 \leq (\underline{C}_N^* Q_N^{-1} \underline{C}_N) \quad (\underline{n}_N^* Q_N \underline{n}_N)$$

By the hypothesis,

$$\underline{C}_N^* Q_N^{-1} \underline{C}_N \leq M^2$$

Therefore for arbitrary $\underline{n}_N$

$$|\underline{n}_N^* \underline{C}_N|^2 \leq M^2 \ \underline{n}_N^* Q_N \underline{n}_N$$

or

$$M^2 Q_N - \underline{C}_N \underline{C}_N^*$$

is non-negative definite. Q.E.D.

To see the geometric interpretation, consider the following simple example.

Let

$$N = 2$$

$$Q_2 = \begin{bmatrix} q_{11} & 0 \\ 0 & q_{22} \end{bmatrix}, \qquad q_{11}, q_{22} > 0$$

$$M = 1$$

$\underline{C}_2$ real

Then

$$\underline{C}_2^* Q_2^{-1} \underline{C}_2 = C_1^2 q_{11}^{-1} + C_2^2 q_{22}^{-1}$$

Therefore

$$\underline{C}_2^* Q_2^{-1} \underline{C}_2 \leq 1$$

describes the set of points (C_1, C_2) inside of the ellipse

$$C_1^2 q_{11}^{-1} + C_2^2 q_{22}^{-1} = 1.$$

In order to show that the method of generating the set of reachable set of states has been given, define the set P_N to be

$$P_N = \left\{ \{C_n\}; \quad \underline{C}_N^* Q_N^{-1} \underline{C}_N \leq M^2 \right\}$$

$$N = 1, 2, \ldots$$

Then, in view of corollary 4.1 and theorem 4.2,

$$C = \bigcap_{N=1}^{\infty} P_N$$

The infinite intersection is simplified because of the following theorem.

THEOREM 4.3 $\quad P_{N+1} \subset P_N$

Proof.

Let $\{C_n\}$ belong to P_{N+1}

Then for every sequence of complex numbers

$$\{\eta_1, \ldots, \eta_{N+1}\}, \quad \left|\sum_{n=1}^{N+1} \eta_n C_n\right| \leq M\sqrt{\int_0^{t_1} \left|\sum_{n=1}^{N+1} \eta_n v_n(t)\right|^2 dt}$$

In particular, it must hold for every sequence of the form $\{\eta_1, \ldots, \eta_N, 0\}$, which implies

$$\left|\sum_{n=1}^{N} \eta_n C_n\right| \leq M\sqrt{\int_0^{t_1} \left|\sum_{n=1}^{N} \eta_n v_n(t)\right|^2 dt}$$

and therefore $\{C_n\}$ belongs to P_N. Q.E.D.

4.2 APPLICATION TO COMPLETE CONTROLLABILITY

Although the preceeding material has dealt with controls which were constrained in norm, there are some applications which can be made to complete controllability if the requirement that M be fixed is dropped.

Let

$$R_M = \left\{ y(x,t_1) \varepsilon L_2(\Omega);\ y(x,t_1) = \sum_{n=1}^{\infty} C_n \psi_n(x), \text{ such that} \right.$$

$$\left. \text{for all } N \quad \underset{\sim N}{C^*} Q_N^{-1} \underset{\sim N}{C} \le M^2 \right\}$$

$$M = 1, 2, \ldots$$

R_M is the reachable set in $L_2(\Omega)$ for given M, and for convenience let this set be defined for positive integer values of M.

Let

$$U_M = \left\{ u \varepsilon L_2(T);\ \left[\int_0^{t_1} |u(t)|^2 dt \right]^{\frac{1}{2}} \le M \right\}$$

Since

$$U = \bigcup_{M=1}^{\infty} U_M$$

$$R(L) = \bigcup_{M=1}^{\infty} R_M$$

Therefore, if the sets $\{R_M\}$ are dense in $L_2(\Omega)$,

$$\overline{\bigcup_{M=1}^{\infty} R_M} = L_2(\Omega)$$

and therefore

$$\overline{R(L)} = L_2(\Omega)$$

and the system is completely controllable.

One should not, however, be led by similar reasoning to the erroneous conclusion that since sets $\{S_N\}$ of the form

$$S_N = \left\{ y(x,t_1) \varepsilon L_2(\Omega); \quad y(x,t_1) = \sum_{n=1}^{N} C_n \psi_n(x) \right\}$$

are dense in $L_2(\Omega)$, if the finite dimensional approximation

$$\dot{y}_n(t) = \lambda_n y_n(t) + b_n u(t)$$

$$y_n(0) = 0$$

$$n = 1, 2, \ldots, N$$

is completely controllable for each N, then the system is completely controllable. That is, the ability of the control to affect each mode does not necessarily imply complete controllability. For example, linear independence of $\{v_1, \ldots, v_N\}$ for each finite N will imply each finite dimensional approximation is completely controllable. This condition is not sufficient to guarantee complete controllability. The additional requirement needed to give complete controllability is that if for given $\{C_1, \ldots, C_N\}$, $u(t)$ satisfies

$$C_n = \int_0^{t_1} v_n(t)u(t)dt, \; n=1, \ldots, N$$

then

$$0 = \int_0^{t_1} v_n(t)u(t)dt \qquad n \geq N+1$$

It is interesting to investigate the possibility of exactly achieving states of the form $\{C_1, C_2, \ldots, C_N, 0, \ldots\}$ for arbitary $\underline{C}_N$ and the remaining elements zero, i.e., $C_n = 0 \quad n \geq N+1$. Suppose that it is desired to attain the state $\{C_1, 0, 0, \ldots\}$ exactly, i.e., the ability to achieve the first mode exactly. The possibility of doing this is contained in the following theorem.

THEOREM 4.4 It is possible to achieve the state

$$C_1 = \text{arbitrary, non-zero}$$

$$C_n = 0 \qquad n = 2, 3, \ldots$$

with a control u such that

$$\int_0^{t_1} u^2(t)dt \leq M^2$$

for some M if and only if

$$\underset{\{\eta_1, \ldots, \eta_N\}}{\text{Inf}} \left|\left|v_1 - \sum_{n=2}^{N} \eta_n v_n\right|\right|_T = L > 0 \tag{4.10}$$

Proof. The necessary and sufficient condition for achieving the specified state is that for every finite sequence $\underline{\eta}_N$,

$$|\underline{\eta}_N^* \underline{C}_N| \leq M \cdot \left|\left| \sum_{n=1}^{N} \eta_n v_n \right|\right|_T \tag{4.11}$$

The left hand side of (4.11) is simply $|\bar{\eta}_1 C_1|$. The inequality holds for $\eta_1 = 0$, therefore it is sufficient to test for $\eta_1 = 1$. The necessary and sufficient condition becomes

$$|C_1| \leq M \,||\, v_1 + \sum_{n=2}^{N} \eta_n v_n \,||_T \qquad (4.12)$$

for arbitrary $\{\eta_2, \ldots, \eta_n\}$. Therefore, if L>0, it is possible to achieve the state $|C_1|$, with a control whose norm is less than or equal to $\frac{|C_1|}{L}$. Conversely, if L = 0, (4.12) requires that $|C_1| = 0$. Q.E.D.

The above theorem implies that if it is possible to expand v_1 in terms the remaining $\{v_2, v_3, \ldots\}$ i.e.,

$$v_1 = \sum_{n=2}^{\infty} \eta_n v_n \qquad (4.13)$$

then it is not possible to achieve the first mode exactly. Note that it is possible that $\{v_1, \ldots, v_N\}$ are linearly independent for all N, but that there is expansion of v_1 given in equation (4.13).

4.3 EXAMPLE

Let the system be described by the wave equation with the control appearing at the boundary in the following form.

$$
\begin{aligned}
\frac{\partial^2 y}{\partial t^2} &= \frac{\partial^2 y}{\partial x^2} \\
\Omega &= (0,1) \\
y(x,0) &= \frac{\partial y(x,0)}{\partial t} = 0 \\
y(0,t) &= 0 \\
\frac{\partial y(1,t)}{\partial x} &= u(t)
\end{aligned}
\tag{4.14}
$$

The control is of the slope at the edge $x = 1$.

The equivalent homogeneous boundary value problem is

$$\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2} + \delta(x-1)u(t) \tag{4.15}$$

$$
\begin{aligned}
y(x,0) &= \frac{\partial y(x,0)}{\partial t} = 0 \\
y(0,t) &= 0 \\
\frac{\partial y(1,t)}{\partial x} &= 0
\end{aligned}
$$

This problem is an example of one in which difficulties arise in a straight forward attempt to reduce it to a first order system by the introduction of a two component vector, v, where

$$v = \begin{bmatrix} y \\ \\ \frac{\partial y}{\partial t} \end{bmatrix}$$

The partial differential equation which v must satisfy is

$$\begin{bmatrix} \frac{\partial y}{\partial t} \\ \\ \frac{\partial^2 y}{\partial t^2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \\ \frac{\partial^2}{\partial x^2} & 0 \end{bmatrix} \begin{bmatrix} y \\ \\ \frac{\partial y}{\partial t} \end{bmatrix} + \begin{bmatrix} 0 \\ \\ \delta\ (x-1) \end{bmatrix} \quad u(t)$$

$$\begin{bmatrix} y(x,0) \\ \\ \frac{\partial y(x,0)}{\partial t} \end{bmatrix} = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix}$$

$$[1,\ 0] \quad \left. \begin{bmatrix} y(x,t) \\ \\ \frac{\partial y(x,t)}{\partial t} \end{bmatrix} \right|_{x=0} = 0$$

$$[\tfrac{\partial}{\partial x},\ 0] \quad \left. \begin{bmatrix} y(x,t) \\ \\ \frac{\partial y(x,t)}{\partial t} \end{bmatrix} \right|_{x=1} = 0$$

The problem is now expressed in the form

$$\frac{\partial v}{\partial t} = Av + F(x)u(t) \tag{4.16}$$

$$v(x,0) = 0$$

$$Uv = 0$$

The difficulty which arises is in the well posedness of the problem by this choice of state variable. Richtmyer [21, Ch. 8] shows that the initial value problem for a system, i.e.,

$$\frac{\partial v}{\partial t} = Av \tag{4.17}$$

$$v(x,0) = v_o(x)$$

$$Uv = 0$$

for some initial state $v_o(x)$ may not be well posed when reduced to a first order system even though the original problem is well posed. Brogan [2] also comments on this situation. For this example, suppose the initial state is

$$v_o(x) = \varepsilon \begin{bmatrix} \sin \omega_N x \\ 0 \end{bmatrix}$$

where

$$\omega_N = (2N - 1) \frac{\pi}{2}$$

The solution is then found to be

$$v(x,t) = \varepsilon \begin{bmatrix} \sin \omega_N x \quad \cos \omega_N t \\ \\ -\omega_N \sin \omega_N x \sin \omega_N t \end{bmatrix}$$

The square of the norm of the initial state is

$$||v_o(x)||^2_\Omega = \varepsilon^2/2$$

whereas at time t

$$||v(x,t)||^2_\Omega = \varepsilon^2/2 \left[\cos^2 \omega_N t + \omega_N^2 \sin^2 \omega_N t \right]$$

Thus, even though the initial state can be made arbitrarily small by the choice of ε, the future states can be arbitrarily large since ω_N is of the order N. Richtmyer [21] shows that the correct choice of the state variable for this system is

$$v(x,t) = \begin{bmatrix} \frac{\partial y}{\partial t} \\ \frac{\partial y}{\partial x} \end{bmatrix}$$

Rather than follow Richtmyer's approach, it will be shown that it is possible to treat equation (4.15) directly and all of the results of this chapter will apply.

Let $\{\lambda_n\}$ and $\{\psi_n\}$ be the eigenvalues and eigenfunctions of

$$\frac{d^2\psi_n(x)}{dx^2} = \lambda_n \psi_n(x), \quad x \varepsilon \Omega \tag{4.18}$$

$$\psi_n(0) = \psi_n'(1) = 0$$

Since the problem is self-adjoint in the scalar valued $L_2(\Omega)$ space, the set $\{\psi_n\}$ is orthonormal and complete [22, Ch. 7]. In this case

$$\psi_n(x) = \sqrt{2}\ \sin \omega_n x \tag{4.19}$$

where

$$\omega_n = (2n - 1)\ \frac{\pi}{2} \tag{4.20}$$

and

$$\lambda_n = -\omega_n^2 \tag{4.21}$$

Any function y in the $L_2(\Omega x T)$ space has an expansion

$$y(x,t) = \sum_{n=1}^{\infty} y_n(t)\psi_n(x)$$

where

$$y_n(t) = \int_0^1 y(x,t)\psi_n(x)dx$$

Substituting the above into equation (4.15)

$$\sum_{n=1}^{\infty} \ddot{y}_n(t)\psi_n(x) = \sum_{n=1}^{\infty} \lambda_n y_n(t)\psi_n(x) + \delta(x-1)u(t) \tag{4.22}$$

Multiplying through equation (4.22) by ψ_n and integrating over (0,1), the countably infinite set of ordinary differential equations results,

$$\ddot{y}_n(t) = \lambda_n y_n(t) + \psi_n(1)u(t) \tag{4.23}$$

where use has been made of

$$\int_0^1 \psi_n(x)\delta(x-1)dx = \psi_n(1)$$

and the orthonormality of the $\{\psi_n\}$.

In order to satisfy the initial conditions,

$$y_n(0) = \dot{y}_n(0) = 0$$

Let

$$z_n(t) = \dot{y}_n(t) + i\omega_n y_n(t) \tag{4.24}$$

Since the partial differential equation (4.14) is second order in t, the state at time t is given by the pair of sequences $\{y_n(t)\}$ and $\{\dot{y}_n(t)\}$. However both of these are contained in the single sequence of complex valued functions $\{z_n(t)\}$. $y_n(t)$ and $\dot{y}_n(t)$ are found by

$$y_n(t) = \frac{z_n(t) - \bar{z}_n(t)}{2i\omega_n} \tag{4.25}$$

$$\dot{y}_n(t) = \frac{z_n(t) + \bar{z}_n(t)}{2} \tag{4.26}$$

$z_n(t)$ satisfies the differential equation

$$\dot{z}_n(t) = \ddot{y}_n(t) + i\omega_n \dot{y}_n(t)$$

By equations (4.21) and (4.23),

$$\dot{z}_n(t) = -\omega_n^2 y_n(t) + \psi_n(1)u(t) + i\omega_n y_n(t)$$

Using equations (4.25) and (4.26),

$$\dot{z}_n(t) = -\omega_n^2 \frac{z_n(t) - \bar{z}_n(t)}{2i\omega_n} + i\omega_n \frac{z_n(t) + \bar{z}_n(t)}{2}$$

$$+ \psi(1)u(t)$$

Combining terms

$$\dot{z}_n(t) = i\omega_n \left[\frac{z_n(t) - \bar{z}_n(t)}{2} + \frac{z_n(t) + \bar{z}_n(t)}{2} \right]$$

$$+ \psi_n(1)u(t)$$

The differential equation for $z_n(t)$ is

$$\dot{z}_n(t) = i\omega_n z_n(t) + \psi(1)u(t)$$

with initial conditions

$$z_n(0) = 0$$

The solution is

$$z_n(t) = \int_0^t e^{i\omega_n(t-\tau)} \psi_n(1)u(\tau)d\tau$$

To put this into the general form studied in this chapter, let the time be fixed at t_1 and let

$$C_n = z_n(t_1)$$

$$v_n(\tau) = e^{i\omega_n(t_1-\tau)}\psi_n(1)$$

Now the necessary and sufficient conditions for $\{C_n\}$ to belong to the set of reachable states is that there exists a control u satisfying

$$\left[\int_0^{t_1} |u(t)|^2\, dt\right]^{\frac{1}{2}} \le M \tag{4.27}$$

and

$$C_n = \int_0^{t_1} v_n(t)u(t)dt \tag{4.28}$$

for all n

As before, define the elements q_{mn} of the matrix Q_N as

$$q_{mn} = \int_0^{t_1} \bar{v}_m(\tau)v_n(\tau)d\tau$$

It will be convenient for this example to pick t_1 to be

$$t_1 = 2$$

then

$$q_{mn} = \psi_m(1)\psi_n(1) \int_0^{2} e^{i(\omega_n - \omega_m)(2-\tau)} d\tau$$

From equation (4.20)

$$\omega_n - \omega_m = (n - m)\pi$$

Therefore

$$e^{2i(\omega_n - \omega_m)} = 1 \qquad \text{for all } m, n$$

When $m \neq n$,

$$\int_0^2 e^{-i(\omega_n - \omega_m)\tau} d\tau = \frac{e^{-i(\omega_n - \omega_m)\tau}\Big|_0^2}{-i(\omega_n - \omega_m)}$$

$$= 0$$

If $m = n$,

$$q_{nn} = 2\,\psi_n^2(1)$$

From equations (4.19) and (4.20),

$$\psi_n(1) = \sqrt{2} \quad \sin\,(2n-1)\frac{\pi}{2}$$

therefore

$$\psi_n^2(1) = 2$$

Thus

$$q_{mn} = 4\,\delta_{mn}$$

The necessary and sufficient conditions on $\{C_n\}$ for a solution to equation (4.28) is, by theorem 4.2, that for all N,

$$\underline{C}_N^* Q_N^{-1} \underline{C}_N \leq M^2 \tag{4.29}$$

In this case,

$$\underline{C}_N^* Q_N^{-1} \underline{C}_N = \frac{1}{4} \sum_{n=1}^{N} |C_n|^2$$

In order that the inequality in (4.29) hold for all N, it is necessary and sufficient that

$$\frac{1}{4} \sum_{n=1}^{\infty} |C_n|^2 \leq M^2 \tag{4.30}$$

Relating the $|C_n|^2$ to $y_n(t_1)$ and $\dot{y}_n(t_1)$

$$|C_n|^2 = |z_n(t_1)|^2$$

From equation (4.24),

$$|z_n(t_1)|^2 = \dot{y}_n^2(t_1) + \omega_n^2 \, y_n^2(t_1)$$

So that the inequality (4.30) becomes

$$\sum_{n=1}^{\infty} [\omega_n^2 \, y_n^2(t_1) + \dot{y}_n^2(t_1)] \leq 4M^2 \tag{4.31}$$

Thus, the above inequality specifies the necessary and sufficient conditions for the sets $\{y_n(t_1)\}$ and $\{\dot{y}_n(t_1)\}$ to belong to the reachable set at time $t_1=2$ when the control is constrained in norm according to (4.27).

If the requirement that M is fixed is dropped, the following result is obtained.

THEOREM 4.5. The system described by Equation (4.14) is completely controllable at time $t_1=2$.

Proof. The state of the system at time t_1 is the pair $y(x,t_1)$, $\frac{\partial y(x,t_1)}{\partial t}$. In order that the system be completely controllable, it is necessary and sufficient that the reachable set of states $y(x,t_1)$ and $\frac{\partial y(x,t_1)}{\partial t}$ be dense in $L_2(\Omega)$. The reachable set of states are those given by

$$y(x,t_1) = \sum_{n=1}^{\infty} y_n(t_1)\psi_n(x) \tag{4.32}$$

$$\frac{\partial y(x,t_1)}{\partial t} = \sum_{n=1}^{\infty} \dot{y}_n(t_1)\psi_n(x). \tag{4.33}$$

When the requirement on u is simply that

$$\int_0^{t_1} |u(t)|^2 \, dt < \infty$$

this requires that for some M,

$$\int_0^{t_1} |u(t)|^2 \, dt \le M^2$$

However, the set of states $\{y(x,t_1), \frac{\partial y(x,t_1)}{\partial t}\}$ satisfying equations (4.32) and (4.33) where $\{y_n(t_1)\}$ and $\{\dot{y}_n(t_1)\}$ satisfy (4.31) for some M are dense in the $L_2(\Omega)$ space.

I.e., let R be the reachable set.

$$R = \left\{ \left(y(x,t_1),\ \frac{\partial y(x,t_1)}{\partial t}\right);\quad y(x,t_1) = \sum_{n=1}^{\infty} y_n(t_1)\psi_n(x) \right.$$

$$\frac{\partial y(x,t_1)}{\partial t} = \sum_{n=1}^{\infty} \dot{y}_n(t_1)\psi_n(x)$$

such that for some M

$$\left. \sum_{n=1}^{\infty} [\omega_n^2 y_n^2(t_1) + \dot{y}_n^2(t_1)] \leq 4\ M^2 \right\}$$

Then R is dense in the product space $L_2(\Omega) \times L_2(\Omega)$. Q.E.D.

Note that R above is not the whole product space $L_2(\Omega) \times L_2(\Omega)$.

In order for R to be the whole space, the requirement on $\{y_n(t_1)\}$ and $\{\dot{y}_n(t_1)\}$ would have to be relaxed to

$$\sum_{n=1}^{\infty} [y_n^2(t_1) + \dot{y}_n^2(t_1)] \leq 4\ M^2$$

for some M. Since ω_n is of the order of n, R is merely a dense subset.

4.4 SUMMARY

The new results in this chapter are the applications of the moment problem to distributed parameter systems. Numerous applications have appeared for finite dimensional control systems. The most closely related have been those by Antosiewicz [9] and Kreindler [11]. Corollary 4.1 of this chapter appears in Antosiewicz [9]. Kreindler [11] states a result similar to Theorem 4.2 and has several examples of reachable sets for finite dimensional systems.

Russell [6] has shown the controllability of the wave equation and the beam equation by methods different from those used in this chapter.

PRECEDING PAGE BLANK NOT FILMED.

BIBLIOGRAPHY

1. Wang, P. K. C. "Control of Distributed Parameter Systems," Advances in Control Systems, 1, (C. T. Leondes, editor), Academic Press, New York, 1964, pp. 75-172.

2. Brogan, W. L. Optimal Control Theory Applied to Systems Described by Partial Differential Equations, Ph. D. Dissertation, Department of Engineering, University of California, Los Angeles, 1965.

3. Fattorini, H. O. "Some Remarks on Complete Controllability," SIAM Journal on Control, 4, (Nov. 1966), pp. 686-694.

4. ______________ "Control in Finite Time of Differential Equations in Banach Space," Comm. on Pure and Applied Math., 19, (1966), pp 17-34.

5. Balakrishnan, A. V. "Optimal Control Problems in Banach Spaces," SIAM Journal on Control, 3, (1965), pp. 152-180.

6. Russell, D. L. "On Boundary Value Controllability of Linear Symmetric Hyperbolic System," Conference on the Mathematical Theory of Control, Los Angeles, 1967.

7. Kalman, R. E., Y. C. Ho, and K. S. Narendra, "Controllability of Linear Dynamical Systems," Contributions to Differential Equations, (1963), pp. 189-213.

8. Hsieh, H. C "On the Final Value and Minimum Effort Systems," Journal of the Franklin Institute, 276, (1963), pp. 154-167.

9. Antosiewicz, H. A. "Linear Control Systems," Archive for Rational Mechanics and Analysis, 12, (1963), pp. 313-324.

10. Banach, S. Theorie des Operations Linearires, Warsaw, 1932.

11. Kreindler, E. "Contributions to the Theory of Time-Optional Control," Journal of the Franklin Institute, 275, (1963), pp. 314-344.

12. Friedman, B. Principles and Techniques of Applied Mathematics, Wiley, New York, 1956.

13. Taylor, A. E. Introduction to Functional Analysis, Wiley, New York, 1958.

14. Penrose, R. "A Generalized Inverse for Matrices," Proc. of the Cambridge Philosphical Society, 51, (1955), pp. 406-413.

15. Loud, W. S. "Generalized Inverses and Generalized Greens' Functions," J. SIAM Appl. Math., 14, (1966), pp. 342-369.

16. Zadeh, L. and C. A. Desoer, Linear Systems Theory, McGraw-Hill Book Co., New York, 1963.

17. Hsieh, H. C. Synthesis of Adaptive Control Systems by Function Space Methods," Ph. D. Dissertation, Department of Engineering, University of California, Los Angeles, 1963.

18. Balakrishnan, A. V. "An Operator Theoretic Formulation of a Class of Control Problems and a Steepest Descent Method of Solution," SIAM Journal on Control, 1, (1963), pp. 109-127.

19. Sakawa, Y. "Solution of an Optimal Control Problem in a Distributed Parameter System," IEEE Transactions on Automatic Control, 9, (1964), pp. 420-426.

20. Yosida, K. Functional Analysis, Academic Press, New York, 1965.

21. Richtmyer, R. D. Difference Methods for Initial - Value Problems, Interscience Publishers, New York, 1957.

22. Coddington, E. A. and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill Book Co., New York, 1955.